

DOCUMENT RESUME

ED 202 909

TM 810 379

AUTHOR Schmitt, Alicia P.; Crocker, Linda
 TITLE Improving Examinee Performance on Multiple Choice Tests.
 PUB DATE Apr 81
 NOTE 15p.; Paper presented at the Annual Meeting of the American Educational Research Association (65th, Los Angeles, CA, April 13-17, 1981).
 EDRS PRICE MF01/PC01 Plus Postage.
 DESCRIPTORS Error of Measurement; Higher Education; *Multiple Choice Tests; *Response Style (Tests); *Test Anxiety; Testing Problems; Test Reliability; Test Wiseness
 IDENTIFIERS Mandler Sarason Test Anxiety Questionnaire

ABSTRACT

The effectiveness of a strategy for improving performance on multiple choice items for examinees with different levels of test anxiety was assessed. Undergraduate measurement students responded to the Mandler-Sarason Test Anxiety Scale and to an objective test covering course content. Results indicated that, for most examinees, generation of an answer before selecting a multiple choice response led to higher test performance, but for highly test anxious examinees, there was a negative effect on performance. Students with less test anxiety might have focused their attention more on the task at hand and thus used more effectively the thought processes involved in generating answers before approaching the multiple choice alternatives. This approach would be most beneficial to students who have real knowledge of the content being tested. Other variables of the test situation, such as impulsive answering, guessing, and other response sets might also be reduced. The experimental strategy resulted in smaller error of measurement and higher test reliability. (Author/GK)

 * Reproductions supplied by EDRS are the best that can be made *
 * from the original document. *

ED202909

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

IMPROVING EXAMINEE PERFORMANCE
ON MULTIPLE CHOICE TESTS

Alicia P. Schmitt

Linda Crocker

University of Florida

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

A. P. Schmitt &
L. Crocker

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

Paper presented at the annual meeting of the American Educational Research Association held in Los Angeles, CA., April 1981.

IMPROVING EXAMINEE PERFORMANCE
ON MULTIPLE CHOICE TESTS

Alicia P. Schmitt

Linda Crocker

University of Florida

Paper presented at the annual meeting of the American Educational Research Association held in Los Angeles CA. April 1961.

ABSTRACT

The effectiveness of a strategy for improving performance on multiple choice items for examinees with different levels of test anxiety was assessed. Undergraduate measurement students responded to the Mandler-Sarason Test Anxiety Scale and to an objective test covering course content. For most examinees, generation of an answer before selecting a multiple choice response led to higher test performance, but for highly test anxious examinees, there was a negative effect on performance. The experimental strategy resulted in smaller error of measurement and higher test reliability.

IMPROVING EXAMINEE PERFORMANCE
ON MULTIPLE CHOICE TESTS

Alicia P. Schmitt

Linda Crocker¹

University of Florida

Most attempts to improve test taking skills have centered on initial reading skills (Lange, 1978), methods for budgeting time during the test, or "tricks" for finding unintended clues among the alternative responses (Sarnacki, 1979). Equivalent time-using and error-avoidance tactics have been proposed in several "how-to" books on test-taking (Huff, 1961, Millman and Pauk, 1969). Although such strategies have been reasonably effective in making examinees more test-wise and perhaps less test-anxious, they do not really encourage the examinee to use his knowledge of content itself.

The most accepted and used definition of test-wiseness is the one offered by Millman, Bishop, and Ebel (1965), who defined test-wiseness as "a subject's capacity to utilize the characteristics and formats of the test and/or the test-taking situation to receive a high score." and added that "Test-wiseness is logically independent of the examinee's knowledge of the subject matter for which the items are supposedly measures." Effective use of knowledge of subject matter or content is not generally considered to be a part of test-wiseness. Yet strategies for improving examinee test performance need not be restricted to efforts to increase test-wiseness." The purpose of this study was to explore the effect of a test-taking strategy which would allow examinees to capitalize on their knowledge of substantive course content.

¹We gratefully acknowledge James Algina for consultation on the analytic methods employed in this study.

The rationale for this test-taking strategy originated from Rowe's (1978, 1973) research on teacher oral questioning techniques in the classroom. Rowe reported that waiting for answers, instead of giving students the correct answer or prompting them with cues right away, can lead to more speculative student thinking, more self-confidence in responses, and an increased number of inferences supported by evidence. There is a striking parallel between teachers oral questioning behavior and an objective testing situation. Rapid questioning and giving an answer or prompting without waiting is similar to a question in multiple choice format. It seemed possible that the thought-generating process could be short-circuited or discouraged by the presence of multiple choice responses just as it is affected by teacher's oral prompting. If the examinee were to read a question, and then generate his own response before selecting among possible correct answers (multiple choices), then test performance might be improved. Time to think and answer the questions could lead to clarification of what was being asked and identification of the concepts involved in the correct answer. After passing through this step the examinee might find it easier to recognize the correct response or at least eliminate irrelevant foils. Some experts (e.g., Millman and Pauk, 1969) have suggested that examinees might benefit from trying to "anticipate the desired response", but they offered no formal strategy accomplishing this and no empirical evidence of its effectiveness.

The objectives of the present study were to (1) assess the effectiveness of response generation in improving performance on multiple choice items; (2) study its effect on examinees when test anxiety level was controlled; and (3) evaluate its influence on error of measurement and test reliability.

METHOD

Subjects

Participants in this study were 86 advanced undergraduate or graduate students enrolled in four sections of an introductory course in tests and measurement.

Instruments

Instruments used in this study were the 37-item Mandler-Sarason Test Anxiety Scale (TAS) and a 15-item objective test covering the first course unit on statistical concepts. Only the first unit of content was selected since subsequent course content on test construction might have affected examinees test-wiseness or test anxiety levels. Two separate test forms were prepared. An experimental test form was designed with instructions and format to assist students in generating their own responses. On this form, each item-stem was presented as an open-ended question with space provided for students to record an answer before turning to the next page to read the alternate responses for that item. The second test form consisted of the same 15 items presented in traditional multiple choice format (control condition).

Procedures for Data Collection and Analysis

The test anxiety measure was administered to all students during the first week of class. Approximately three weeks later, the objective test over course material was administered. Examinees were randomly assigned to either the response-generating, multiple choice condition or to the traditional multiple choice condition. No other instructions about test-taking were provided.

Data were analyzed in a oneway analyses of covariance. The dependent variable was the number of correct responses chosen for the multiple choice items; the independent variable with two levels was test form, test anxiety score was the covariate in this analyses. Cases with missing observations were deleted and the

data set was tested for statistical outliers using Cook's (1977) distance criterion. Included in the analyses was a test of the assumption of no covariate treatment interaction.

RESULTS

The means, standard deviations and number of subjects for each group on the covariate test anxiety (TAS) and the classroom achievement test are presented in Table 1. Results of the analyses of covariance are presented in Table 2. Thirteen cases were deleted from the data set due to missing observations and one case was deleted after being identified as an extreme outlier using Cook's (1977) distance criterion. The test for homogeneity of slopes for the regression of the covariate (test anxiety) on the dependent variable (achievement test score) for the two treatment groups (see Table 2) indicated a significant disordinal interaction between the covariate and the treatment. ($F(1, 68) = 5.34, p < .05$). Figure 1 depicts this relationship. The presence of such a significant interaction precluded interpretation of the test of overall treatment effect, but it also indicated that there were differential treatment effects for subjects at certain levels of test anxiety.

The Johnson-Neyman technique was employed to establish regions of significance for the interaction (Kerlinger, and Pedhazur, 1973). This technique indicated how high (and how low) test anxiety scores should be in order for treatment group means to be significantly different. Results of this follow-up analysis indicated that students with a TAS-score lower than 17.72 performed significantly better on their achievement tests under the response-generating experimental condition. (See Table 3) For those students with TAS-scores higher than 21.50 the control group performed significantly better. Dotted lines on Figure 1 delineate the regions of significance. There is a narrow region of approximately four points in the TAS distribution for which there was no discernible treatment effect.

Students with TAS-scores between 17.72 and 21.50 did not differ significantly in their achievement test performance. A majority of the subjects had TAS scores in the lower region of significance (37 out of 72 or 51%). Only seventeen, or 24%, fell on the upper region of significance and 18 or 25% of scores fell in the region of no significant difference.

It should be noted that for the experimental group the correlation between test scores on test anxiety was $-.57$ accounting for 32% of observed variance in achievement test scores. For the control group the relationship was also negative ($r = -.22$) but accounted for only 5% of the variation in test performance.

Internal consistency estimates (Kuder-Richardson 20) and standard error of measurement for the achievement test were computed for each group and total sample. KR20 coefficients were $.43$ for the control group and $.73$ for the experimental group. The KR20 estimate for the total group was $.62$. Standard error of measurement for the control group test was 1.55 ; for the experimental group's test, 1.44 , and for the total group, 1.51 .

DISCUSSION

Results indicated that generation of an answer before trying to select a correct response, improves performance for students with average and below-average levels of test-anxiety (TAS scores ≤ 18). In the present sample most of the subjects (51%) fell in this group. For these students there seemed to be real value in formulating an answer to avoid confusion when confronted with foils. For students with high levels of test-anxiety (TAS scores ≥ 22) the regular multiple choice presentation of items produced better test scores. Twenty-four percent of the sample were in this category. Wine (1971) has suggested an attentional interpretation of the debilitating effects of test-anxiety. Persons with little test-anxiety attend mostly to task-relevant variables, while highly test-anxious persons attend to both self-relevant and task-relevant variables. This might explain the interaction found between TAS and treatment. Students with less

test anxiety might have focused their attention more on the task at hand and thus used more effectively the thought processes involved in generating answers before approaching the multiple choice alternatives.

In the cases where students were highly test-anxious, self-preoccupation might have obstructed the positive effects of the experimental condition. Generating an answer required them to engage in a longer task which might have, indeed, increased their self-preoccupation and lowered their performance. More research is needed in order to explain what happens to highly test-anxious students under response generation. According to Tryon (1980) test-anxiety has been repeatedly correlated with aptitude and achievement tests and grade point average. Significant correlations in these studies range from $-.47$ to $-.14$. In the present study the relationship between TAS and test scores for the experimental group was $-.57$. This indicates that the relationship between these two variables may be dependent on the testing situation and points to the possibility of considering other situational and personal variables when interpreting TAS correlates.

Results indicate that answer generation before approaching foils in multiple choice items can be easily learned and may help improve performance in multiple choice tests. This approach would be most beneficial to students who have real knowledge of the content being tested. Other variables of the test situation, such as impulsive answering, guessing, and other response sets might also be reduced.

One caution to over-zealous interpretation of these results should be noted. In this study, students were actually required to write out their response before seeing the alternate choices and were allowed sufficient time to complete the task. In speeded testing situations, examinees would need to form their answers mentally before examining the choices. It is not yet known whether similar results would be obtained under those circumstances, but further study of the usefulness of response-generation seems warranted.

REFERENCES

- Cook, R. D. - Detection of Influential Observation in Linear Regression. Technometrics. 1977. 19(1), 15-18.
- Huff, D. - Score: The Strategy of Taking Tests. New York: Appleton-Century-Crofts, 1961.
- Kerlinger, F. N., Pedhazur, E. J. - Multiple Regression in Behavioral Research. Holt, Rinehart and Winston, Inc., New York, 1973.
- Lange, R. - Flipping the Coin: From Test Anxiety to Test-Wiseness. Journal of Reading, 1978, 22(3), 275-277.
- Millman, J.; Bishop, C. H.; & Ebel, R. - An Analysis of Test-Wiseness. Educational and Psychological Measurement, 1965, 25, 707-726.
- Millman, J.; & Pauk, W. - How to Take Tests. New York: McGraw-Hill. 1969.
- Rowe, M. B. - Wait, Wait, Wait... School Science and Mathematics, 1978, 78, 207-16.
- Rowe, M. B. (Ed) - What Research Says to the Science Teacher. National Science Teachers Assoc., Washington, D.C.; 1973.
- Sarnacki, R. E. - An Examination of Test-Wiseness in the Cognitive Test Domain. Review of Educational Research, 1979, 49(2), 252-279.
- Tryon, G. S. - The Measurement and Treatment of Test Anxiety. Review of Educational Research, 1980, 50(2), 343-372.
- Wine, J. - Test Anxiety and Direction of Attention. Psychological Bulletin, 1971, 76, 92-104.

TABLE 1

Means and Standard Deviations
by Treatment Group and Total Sample

	Control	Experimental	Total
Test Anxiety Scale (TAS) (37 items)			
N	34	38	83
$\bar{X}$	15.24	13.16	16.42
SD	6.71	6.40	6.46
Achievement Test (15 items)			
N	36	38	74
$\bar{X}$	10.06	10.37	10.22
SD	2.08	2.74	2.43

TABLE 2

Summary of Covariance Analysis
for Effects of Test Form on Achievement Score

Source	df	SS	MS	F
TAS	1	67.70	67.70	15.84
Treatment	1	29.00	29.00	6.78
Interaction	1	24.95	24.95	5.84*
Error	68	290.71	4.28	

* Significant at $\alpha \leq .05$.

TABLE 2

Means and Standard Deviations
of Achievement Test Scores for Different Regions
of Test Anxiety

	Control	Experimental
REGION:		
TAS $\leq$ 17		
N	22	15
$\bar{X}$	10.32	12.2
SD	1.86	1.15
REGION:		
TAS = 18-21		
N	6	12
$\bar{X}$	10.67	10.58
SD	1.86	1.83
REGION:		
TAS $\geq$ 22		
N	6	11
$\bar{X}$	10	7.64
SD	1.79	2.98

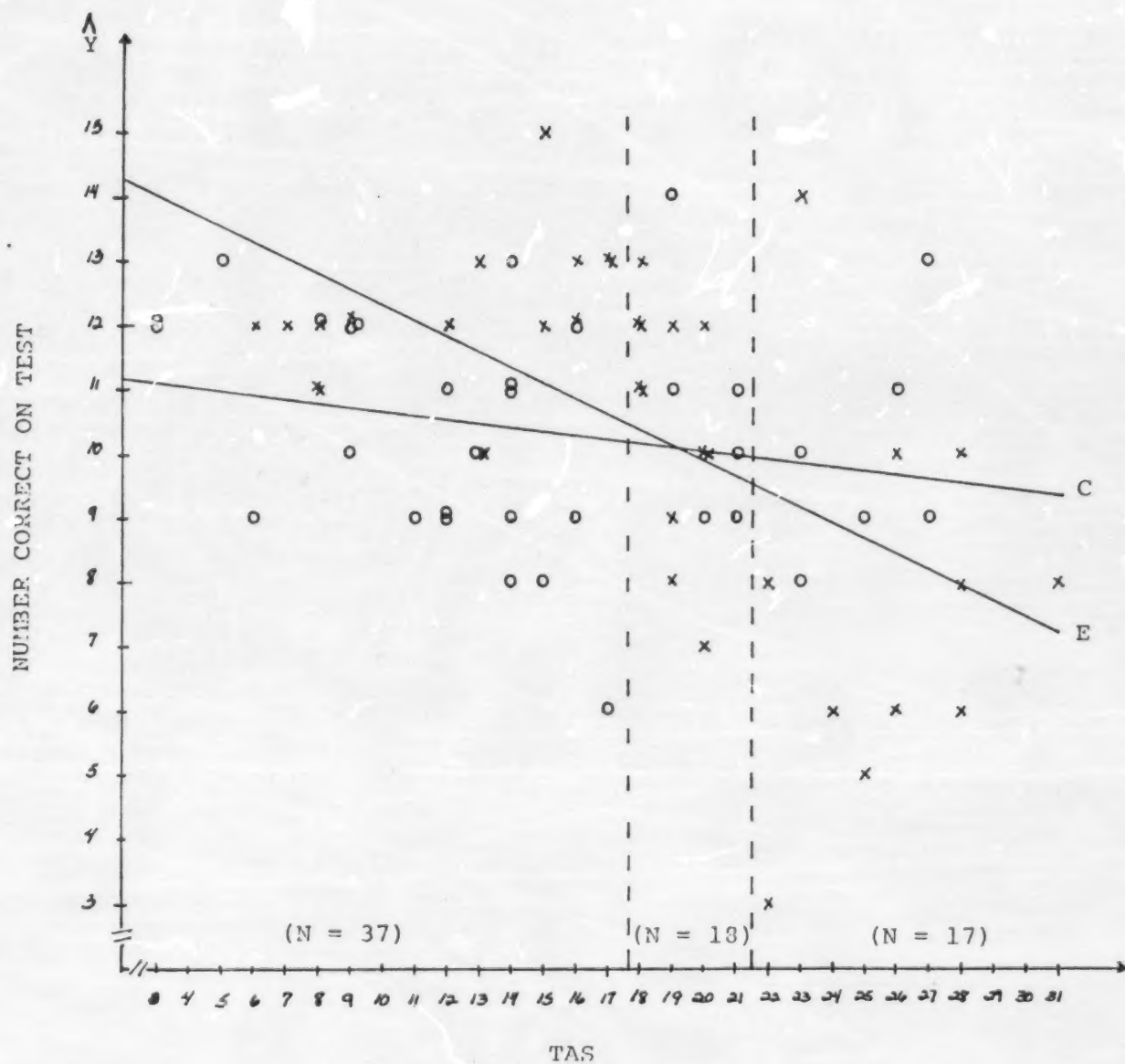


FIGURE 1. Treatment group by test anxiety interaction